

中国海外投资清单（COFI）数据库技术说明

CHINA OVERSEAS FINANCE INVENTORY DATABASE

周李焕，马子轶，刘爽，马欣悦，KEVIN ACKER，MARGARET MYERS，DEBORAH BRAUTIGAM，KEVIN GALLAGHER

摘要

本技术说明解释了中国海外投资清单（COFI）数据库的建立过程。中国海外投资清单是一个涵盖中国在“一带一路”倡议国家发电领域的股权和债权投资的综合数据库。本技术说明阐述了作者如何在公开数据和商业数据库中匹配电厂及其投资，以及如何将最终整合而成的数据库作为开放数据源发布。

引言

中国领导层越来越意识到环境可持续发展是“一带一路”倡议成功的重要因素。“一带一路”倡议由中国国家主席习近平在2013年首次公开提出，致力于推动跨区域发展与经济合作（外交部，2015），交通和能源基础设施是“一带一路”倡议的重点领域。“一带一路”倡议强调在基础设施建设中考虑环境和气候变化的影响。生态环境部发起成立了“一带一路”绿色发展国际联盟，旨在促进“一带一路”绿色发展（BRIGC，n.d.）。追踪中国在“一带一路”的投资数据对于了解中国“一带一路”投资的绿色程度至关重要。更加全面的数据也有助于中国的政策制定者、金融机构和其他利益相关方履行责任，确保“一带一路”绿色发展。

尽管中国的海外基础设施投资有着重大的影响，但目前还没有关于这些交易和项目的全面的数据库。中国政府尚未披露交易或项目层面投资的官方统计数据。现有的公开数据库多以部门、地域和投资者为关注点，且缺乏足够的详细信息，关于商业银行贷款和股权投资的公开数据覆盖面很小。此外，不同来源的数据往往格式各不相同，需要加以统一。因此，应建立一个全面的数据库，为分析和决策提供支持。

目录

摘要	1
1. 引言	1
2. 范围	2
3. 数据来源	26
4. 整合方法	26
5. 结果	12
6. 数据库使用和维护计划	15
附录一：截至2021年1月的“一带一路”国家列表	16
附录二：变量列表与描述	18
附录三：详细的自动匹配方案	21
缩略语表	26
参考文献	27
致谢	28

“技术论文”阐述出版物、交互式应用程序和工具的研究或分析方法。

引用建议：周李焕，马子轶，刘爽，马欣悦，Kevin Acker, Margaret Myers, Deborah Brautigam, Kevin Gallagher. 2022. “中国海外投资清单（COFI）数据库技术说明”技术论文，北京：世界资源研究所. <https://www.wri.org/research/china-overseas-finance-inventory-database>

中国海外投资清单（COFI）数据库（以下或简称“数据库”）是一个利用现有公共和商业数据来源和方法建立的中国基础设施领域对外投资综合数据库。该项目由波士顿大学全球发展政策研究中心、美洲对话（Inter-American Dialogue）、约翰霍普金斯大学中非研究所和世界资源研究所（WRI）合作完成。

范围

本数据库中，“投资”是指中国实体向海外实体提供的股权或债权投资。现版本数据库侧重于发电项目，包括电厂的技术特征和财务交易两大类信息，每种类型信息包括多个数据点。电力部门是接收来自中国公司和投资者投资最多的基础设施部门之一，被率先选作数据库的重点涵盖内容（Zhou等，2018）。电力部门也是低碳转型最重要的部门之一，一些技术具有很强的长期碳锁定效应，使新建电厂在化石燃料基础设施的生命周期内持续排放温室气体数十年之久（Sato等，2021）。因此，电力部门实现“去碳化”至关重要。未来，数据库计划继续扩大，逐步涵盖其他能源部门、交通部门等碳密集型基础设施部门。

本数据库中包括以下关键数据点：

■ 金融交易

- 金融工具：股权或债权
- 投资者名称
- 金额
- 融资协议签订年份

■ 技术特征

- 名称
- 装机容量
- 投产年份
- 国家
- 主要燃料类别

针对金融交易，数据库涵盖了来自企业和银行的股权和债权投资。这是发电项目（特别是化石燃料发电项目）最主要的融资形式。中国官方数据证实，“一带一路”能源投资几乎全部来自国有企业，他们通过对外直接投资或政策性银行和商业银行的贷款支持，在“一带一路”能源领域投资（王，2021；国际能源署，2020）。股权投资包括绿地投资和跨国并购（M&A）。其中，绿地投资指在国外建立新的实体，跨国并购则涉及对现有公司的收购。债权投资或贷款包括中国部分银行的投资，如国家开发银行（CDB）、中国进出口银行（China Eximbank）和其他一些商业银行等。

本数据库是一个不断发展和更新的数据库，目前涵盖了2000年至2020年间中国在“一带一路”国家电力部门的投资数据。“一带一路”国家是指与中国签署了“一带一路”谅解备忘录的国家，

截至2021年1月，共有140个国家，参见附录一（“一带一路”门户网站，2021）。为给未来研究提供更多的支持，数据库也包括了2013年首次公开提出“一带一路”倡议之前的投资项目。

数据库在初始阶段主要关注项目融资。项目融资是发展中国家基础设施融资最常见和最重要的方式之一。项目融资一般通过特殊目的公司进行贷款，还款依赖于项目现金流，发起人承担风险较小，更适合于基础设施等资本密集型行业的融资（Dentons，2013）。项目融资不包括资产负债表融资、公司债券、担保债务凭证、中国人民银行中央银行互换额度等形式的投资。这些类型的融资通常较独立项目融资而言透明度更低，相关信息有限（Mawutor，2015）。

数据库整合了现有几个涉及中国不同方面融资信息的源数据库的数据，并不自行收集新的数据。整合后的数据由源数据库的条目互相匹配而成，可以呈现中国投资者参与的发电项目的投资额和类型。数据库整合过程中删除了重复的数据，避免重复计算并解决信息冲突问题。

本技术说明内容结构如下：第三节说明了使用的源数据库；第四节概述了整合方法；第五节展示了数据库整合结果；第六节讨论了数据库的使用和维护计划。附录一列出了截至2021年1月的“一带一路”国家，附录二提供了数据库的详细变量列表和解释，附录三详细描述了数据匹配方法。

数据来源

投资数据

通过对现有的公开数据库和商业数据库的研究发现，将公开数据库和商业数据库结合可以更好地覆盖中国的海外投资。数据库的开发工作采用以下标准，选定四个公开数据库和三个商业数据库作为数据来源（表1）。

- 具备项目或交易级数据。
- 数据收集和验证过程的方法论健全且透明。
- 数据涵盖2000—2020年期间的大部分年份，且最好定期更新。
- 商业数据库提供商允许以合理的价格在公开领域分享数据。

使用来自商业数据库的数据，填补了公开数据库的两方面数据空白。一是使用来自路孚特（Refinitiv Eikon）的贷款数据，弥补了中国商业银行贷款有关数据的缺失。二是使用《金融时报》旗下fDi Markets的对外直接投资绿地投资和路孚特的并购数据，填补股权投资数据的缺口。图1展示了数据库包含的源数据库的相关信息。

表 1 | 本数据库包含的源数据库概览

源数据库名称	所属机构/公司	描述	是否公开	数据颗粒度	涵盖时间段	计入状态	数据收集期	覆盖地理范围	观测量（个）	缩写
中国全球能源融资数据库	波士顿大学全球发展政策研究中心	追踪中国两家政策性银行（国家开发银行和中国进出口银行）提供的全球能源融资数据库	是	项目层面	2000—2020 年	签署贷款合同	持续更新	全球	289	BU-CGEF
中国对非洲贷款数据库	约翰霍普金斯大学中非研究所，自 2021 年 3 月起由波士顿大学管理	收集中国对非洲国家政府和非洲国有企业的贷款数据	是	各阶段项目数据 / 各分级贷款数据	2000—2019 年	签署贷款合同	持续更新	非洲	1077	SAIS-CLA
中国—拉丁美洲融资数据库	美洲对话和波士顿大学全球发展政策研究中心	追踪国家开发银行和中国进出口银行向拉丁美洲国家政府及其国有企业提供融资的情况	是	项目层面	2005—2020 年	签署贷款合同	持续更新	拉丁美洲	94	IAD-GEGI
中国海外电力资产数据库	数据来自一篇原创论文。波士顿大学将其开发为在线工具	追踪中国对发电项目的对外直接投资（包括绿地投资和并购）	是	电厂单元	1893—2026 年	电厂（预计）调试时间	2017 年，每两年更新一次	全球	976	Journal-CGP
路孚特跨国并购数据库（Refinitiv Eikon M&A）	路孚特	收集并购交易数据	否	交易层面	2000—2020 年	交易公告	持续更新	全球	189	Refinitiv-M&A
路孚特贷款数据库（Refinitiv Eikon Loan）	路孚特	收集银行贷款数据	否	交易层面	2000—2020 年	签署贷款合同	持续更新	全球	186	Refinitiv-Loan
fDi Markets	金融时报	收集对外投资绿地投资数据	否	交易层面	2004—2019 年	交易公告	持续更新	全球	202	fDi Markets

注：源数据库中包括两种主要类型的里程碑日期，即交易公告和贷款合同签署，以及电厂（预计）调试年份。对于绿地项目，交易公告通常发生在项目开发早期阶段，而贷款合同签署通常是在一系列谈判结束后的项目开发尾声阶段（Dentons, 2013；DOE, 2015）。并购交易的目标公司通常是已经投入运营的公司。电厂调试是运营阶段的开始，发生在项目开发和建设阶段之后（DOE, 2015）。

来源：作者基于 CARI 等，2021；金融时报，2021；Gallagher, 2021；Gallagher 和 Myers, 2021；Li 等，2020；Refinitiv, 2021

图 1 | 数据库所包括的涵盖不同类型股权和债权投资的源数据库

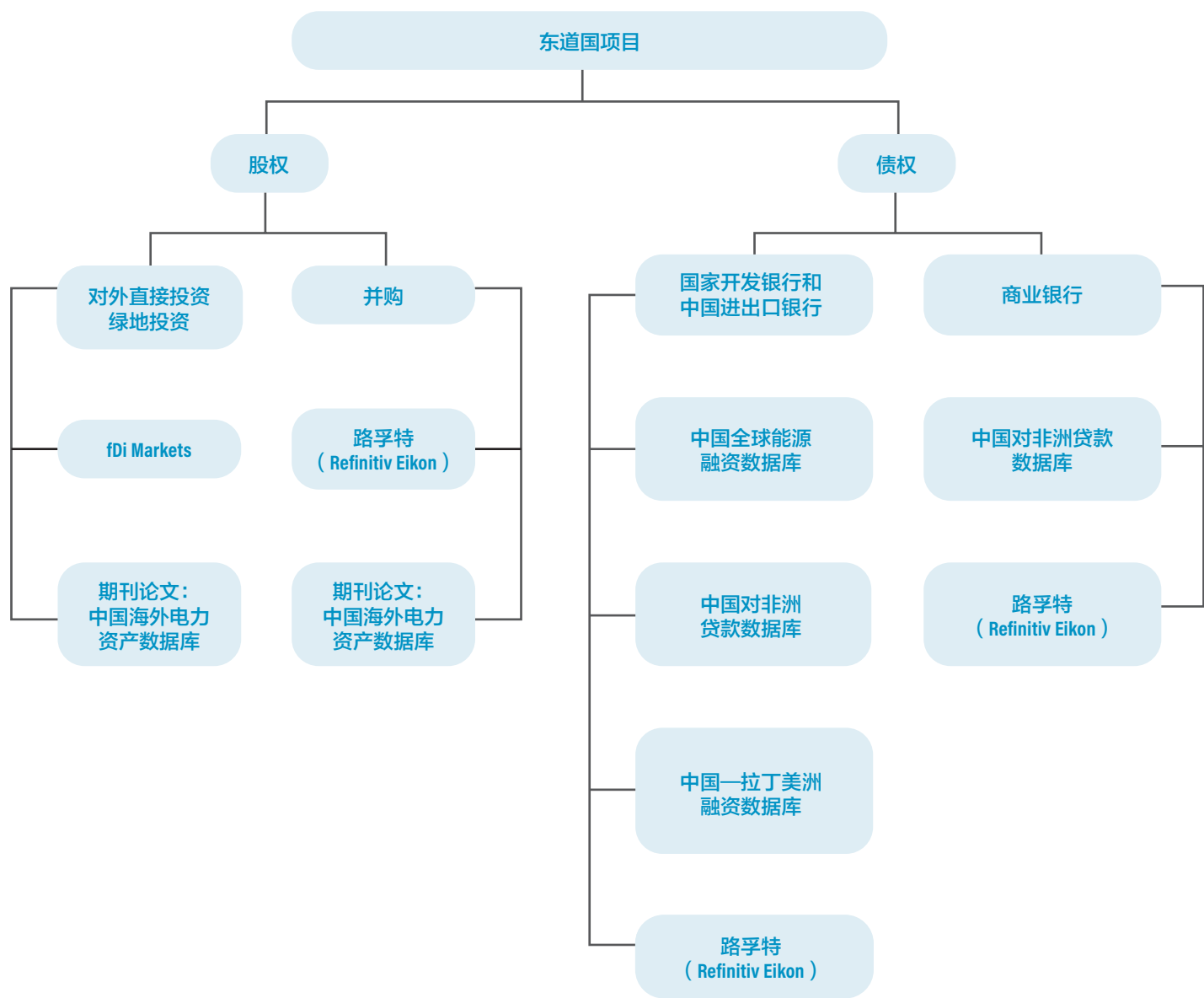


表 2 | 电厂源数据库相关信息

源数据库名称	观测层级	调试年份	数据收集期	电厂地理覆盖范围	覆盖类型	观测数量	观测单位	缩写
UDI 世界电厂数据库（WEPP）	电厂发电单元 ^a	1878—2030 年 ^b	持续更新	全球	全球电厂	211829	电厂发电单元	WEPP
全球电厂数据库（GPPD）	电厂	1896—2018 年	于 2018 年停止更新	全球	全球电厂	29910	电厂 / 电厂发电单元	GPPD

注：^a一些发电单元层面的观测在数据库整合中合并到了电厂层面。附录二对此进行了详细的讨论。
^b由于普氏数据还包括处于规划阶段或在建阶段的电厂，这些电厂的调试年份为估计值，因此调试年份涉及 2018 年之后年份。
来源：作者基于 Byers 等，2018；普氏，2018

电厂特征数据

电厂特征数据依赖于两个源数据库，即普氏（Platts）UDI 世界电厂数据库（WEPP）和全球电厂数据库（GPPD）（Byers 等，2018；普氏，2018）。WEPP和GPPD对投资类源数据库做了补充，提供了地理位置、调试年份、装机容量、主要燃料和其他技术信息（普氏，2018；全球能源观察，2018）。表2提供了数据库开发使用的两个电厂源数据库的相关信息。

覆盖范围的限制

由于本数据库不产生新的数据，数据来源依赖前面提到的源数据库，因此在覆盖面和一些数据条目的准确性方面存在局限。

源数据库中的数据空白包括缺少全部的股权投资额和部分电厂装机容量信息。三个源数据库（fDi Markets、Journal-CGP和Refinitiv-M&A）在股权投资额方面都缺失信息。由于收集支付的股权金额和估计并购交易中的金额比较困难，Journal-CGP和Refinitiv-M&A没有提供这一信息。FDi Markets提供了对外直接投资绿地投资项目的总资本支出，但非中国投资者的股权投资额。

在所有源数据库都提供了一个项目的必要信息的情况下，也存在数据不准确的问题。第一种情况是源数据库未能及时对项目具体情况的变化进行更新。例如，融资结构和所有权在项目的建设和运营过程中可能会发生变化，而这些变化很少被报告，也很难跟踪。因此，源数据库可能无法提供最新的信息。第二种情况是银团贷款结构不明确。有时个别银行的确切贷款额无法从源数据库中获得，有时牵头银行的投资占比可能比其他银行大很多。对于这样的情况，数据库采取的是取总贷款额的平均数，这将在投资方的贷款额估算章节详细讨论。

尽管各个源数据库都致力于展示中国“一带一路”投资的全貌，但源数据仍可能遗漏某些信息。这些源数据库大多依赖二手信息，包括新闻报道和政府网站的信息，而没有中国金融机构和股权投资者直接地、系统地披露的信息。这些二手信息几乎不可能涵盖中国投资的所有项目。例如，整合的数据库希望涵盖中国股权投资基金在“一带一路”国家的投资，如尝试通过几个源数据库涵盖丝路基金的投资信息。周等人（2018年）的研究表明，丝路基金也投资了发电项目，但目前尚无源数据库包含这类来自股权基金的项目投资信息。

整合方法

源数据库通过匹配，可以共同提供一个项目的综合情况。然而，源数据库并不能简单地合并在一起，因为有些项目可能重复出现。通过对来自不同源数据库的条目进行匹配，实现整合，可以避免重复计算的问题，同时，通过整合多个源数据库收集的不同信息，也可以提高数据覆盖的质量。

各种源数据库使用的数据收集方法、分类、定义，以及质量控制标准都不同，给数据整合带来了挑战，也导致数据质量参差不齐。有些源数据库依靠的是新闻报道和政府官方公告，有些则依靠金融机构提供的信息。表3描述了九个源数据库的原始数据来源。不同的源数据库之间也存在数据不一致的情况，包括同一电厂有不同的名称、对观测年份有不同的定义等。例如，一些数源数据库将观测年份定义为电厂的上网年份，而另一些源数据库则将观测年份定义为投资交易的签署或结束年份。此外，一些源数据库中缺少投资额数据。

表 3 | 九个源数据库的原始数据来源

源数据库名称	数据来源
UDI 世界电厂数据库 (WEPP)	直接调查、供应商参考清单、电力公司财务和统计报告，以及贸易和商业新闻
全球电厂数据库 (GPPD)	国家政府机构、电厂建造公司或部件提供公司的报告、公共事业单位数据，以及来自跨国组织的信息
中国全球能源融资数据库	国家开发银行和中国进出口银行的官方网站或东道国的相关部委、新闻报道和官方文件
中国对非洲贷款数据库	借款国政府、银行、公司（包括向美国证券交易委员会提交的文件），以及中国和借款国的新闻报道，当地联系渠道和实地考察
中国—拉丁美洲融资数据库	国家开发银行和中国进出口银行的官方网站或东道国的部委、新闻报道和官方文件
中国海外电力资产数据库	UDI 世界电厂数据库、政府网站、公司网站和年度报告、贸易和商业媒体
路孚特 (Refinitiv Eikon) 跨国并购数据库	来自全球银行业和法律界人士的直接交易资料、监管文件、公司声明、媒体和电汇
路孚特 (Refinitiv Eikon) 贷款数据库	来自全球银行业和法律界人士的直接交易资料、监管文件、公司声明、媒体和电汇
fDi Markets	公开来源：金融时报、媒体、机构、购买数据

来源：作者根据 Byers 等，2018 年；中国非洲研究计划，2021 年；金融时报，2021 年；Gallagher，2021 年；Gallagher 和 Myers，2021 年；李等，2020 年；普氏，2018 年；路孚特，2021 年。

数据整合工作分为三个主要步骤，将源数据库整合为统一而全面的数据库。首先是对数据进行预处理使其标准化，然后通过计算机程序自动匹配各源数据库，最后通过人工匹配确保数据整合的准确性，并解决每个条目的冲突信息。Sonoma技术公司作为合作方，参与了标准化处理和自动匹配工作。本节接下来将详细介绍这三个步骤的工作。

数据预处理

预处理通过三个步骤规范和准备数据，以获得更好的匹配结果。第一步是根据以下标准，筛选相关项目：

- 项目在2000年至2020年间签署贷款合同。例如，签署谅解备忘录，但需要在此初步协议基础上进一步谈判的项目被排除在外（Brautigam和Hwang，2020）。
- 以连接到电网为目的的发电项目。例如，输配电项目和自备电厂被排除在外。
- 项目有来自中国的债权或股权投资。例如，由中国公司承担工程、采购和建设但没有提供融资的项目被排除在外。

- 排除没有信息来源、没有项目细节、无法通过额外研究核实的项目。
- 排除每个源数据库内和源数据库间的重复项目。

第二步是变量名称标准化。这需要对所有源数据库的变量名称进行对应和标准化，以便进行匹配。表4显示了不同源数据库的通用变量，在统一的变量名称行下指定了相应的变量名称。一些源数据库没有在独立栏中提供与这些变量相关的信息，而是一并在“描述栏”中给出。为了尽量利用每个源数据库的所有可用信息，预处理过程将所有可用数据都提取了出来。表5显示了在匹配过程中直接使用的变量以及提取自哪个源数据库。

鉴于不同的源数据库使用了不同的方法和分类来记录基本类别，数据预处理的最后一步是对三个分类变量的类别进行标准化。这三个变量是主要燃料、国家 (iso3c) 和股权投资者名称。其中，燃料来源使用了与数据库一致的标准名称。国家名称通过国家代码包，根据国际标准化组织《国家代码的国际标准》(ISO 3166-1) 中定义的三字母国家代码进行规范 (Arel-Bundock等，2020)。此外，在这一步增加了对投资进行分类的变量。每个条目可以分类为“股权投资”、“债权投资”或“二者兼具”。

表 4 | 源数据库对应的变量名标准化^a

源数据库 变量名	BU-CGEF	SAIS-CLA	FDI MARKETS	IAD-GEGI	JOURNAL-CGP	WEPP	REFINITIV-LOAN	REFINITIV-M&A	GPPD
power_plant_name	description	purpose		purpose		plant	use_of_proceeds_notes	synopsis	name
country	country	country	destination_country	country	country	country	issuer_borrower_nation	target_nation	country_long
country_iso3c	country_code								country
province	location	destination_state				state			
city	location	destination_city							
installed_capacity	mw_if_applicable	purpose	description		capacity_mw	mw	use_of_proceeds_notes		capacity_mw
commissioning_year			description		online_year	year			commissioning_year
primary_fuel	energy_source	purpose	fuel_source		technology	fuel	use_of_proceeds_notes	synopsis	primary_fuel
latitude									latitude
longitude									longitude
total_investment_amountb	amount_m	usd_m	capital_investment	amount					
equity_investor_name_1			investing_company		major_chinese_shareholder	company	acquirer_name		
equity_investor_amount_1			capital_investment					deal_size_m_usd	
equity_investment_type					type_of_investment			form_of_the_transaction	
equity_investment_year			project_date					date_effective	

表 4 | 源数据库对应的变量名标准化^a (续)

源数据库 变量名	BU-CGEF	SAIS-CLA	FDI MARKETS	IAD-GEGI	JOURNAL-CGP	WEPP	REFINITIV-LOAN	REFINITIV- M&A	GPPD
parent_ company_of_ investor				parent_ company		parent		acquirer_ ultimate_ parent	
debt_ investment_year	date	year		year			loan_dates_ facility_closing_ date		
debt_ investment_ amount	amount	us\$		amount					
number_of_ lenders	lender	lender		lender			all_managers		
china_ development_ bank	lender	lender		lender			loan_manager_ commitment_ amount		
export_import_ bank_of_china	lender	lender		lender			loan_manager_ commitment_ amount		
industrial_and_ commercial_ bank_of_china		lender		lender			loan_manager_ commitment_ amount		
china_ construction_ bank		lender					loan_manager_ commitment_ amount		
bank_of_china		lender		lender			loan_manager_ commitment_ amount		
agricultural_ bank_of_china		lender					loan_manager_ commitment_ amount		
china_citic_ bank_corp				lender			loan_manager_ commitment_ amount		
china_ merchants_bank							loan_manager_ commitment_ amount		

表 4 | 源数据库对应的变量名标准化^a (续)

源数据库 变量名	BU-CGEF	SAIS-CLA	FDI MARKETS	IAD-GEGI	JOURNAL-CGP	WEPP	REFINITIV-LOAN	REFINITIV-M&A	GPPD
china_minsheng_banking_corp							loan_manager_commitment_amount		
china_zheshang_bank							loan_manager_commitment_amount		
china_everbright_bank							loan_manager_commitment_amount		
cross-checking ids	bu_id	sais_id	fdi_id	iad_id	unit_id in platts	comp_id; location_id; unit_id	r_id	rma_id	gppd_id

注：a 为方便匹配，所有的变量名都被转换为小写。b 总投资额衡量了电厂的总投资成本。
来源：作者整理

表 5 | 从相应源数据库中提取的变量名称

变量名称	源数据库名称
power_plant_name	BU-CGEF，SAIS-CLA，Refinitiv-M&A，Refinitiv-Loan
capacity	SAIS-CLA，fDi Markets，Refinitiv-Loan
primary_fuel	SAIS-CLA，Refinitiv-M&A，Refinitiv-Loan

来源：作者整理

自动匹配

匹配源数据库的主要挑战是大多数源数据库缺乏共同且唯一的标识，导致无法明确匹配这些记录。数据整合过程使用了R studio软件（R Core Team，2020），两两逐步匹配数据库。图2显示了匹配步骤，附录三则给出了每个步骤的详细匹配计划。

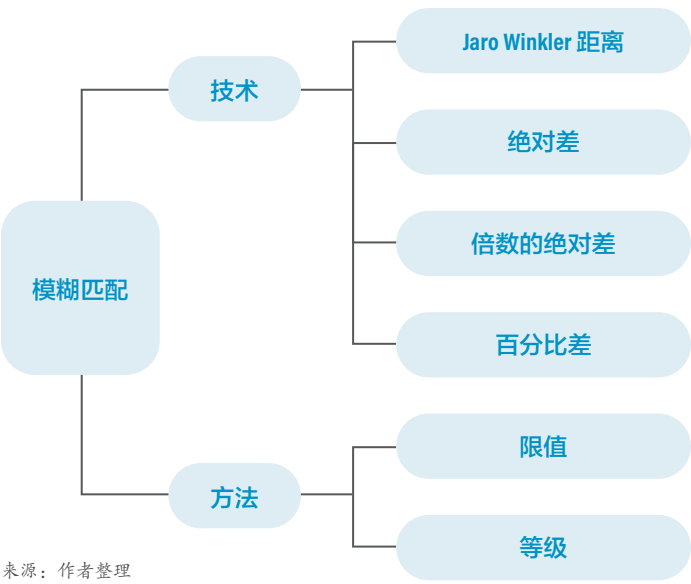
我们采取等值匹配的方法，即在变量等值的基础上匹配字符串，如果源数据库和变量有共同的唯一标识符，则可以匹配它们。这适用于前三个源数据库，即WEPP、Journal-CGP和GPPD。我们使用唯一的标识符UNITID在发电单位层面匹配WEPP和Journal-CGP，然后将匹配结果转换到电厂层面，使用唯一的标识符LOCATIONID进一步与GPPD匹配。UNITID是电厂单位的识别号，LOCATIONID是电厂的识别号。我们还对两个变量country_iso3c（ISO 3166-1 alpha-3国家代码）和primary_fuel使用了等值匹配。通过对这两个变量进行跨数据库的标准化处理，可以检查两个源数据库的数值是否相等。

对于其余的源数据库，我们使用了模糊匹配这种识别非精确匹配的技术，根据共享变量的相似性，将源数据库匹配在一起（图3）（Robinson等，2020）。

我们使用“Jaro-Winkler距离”比较电厂名称等文本之间的相似性。距离越短，则说明两个字符串相似性越高。将“Jaro-Winkler距离”反转，取其与1之间的差值，可得到“Jaro-Winkler分数”（即

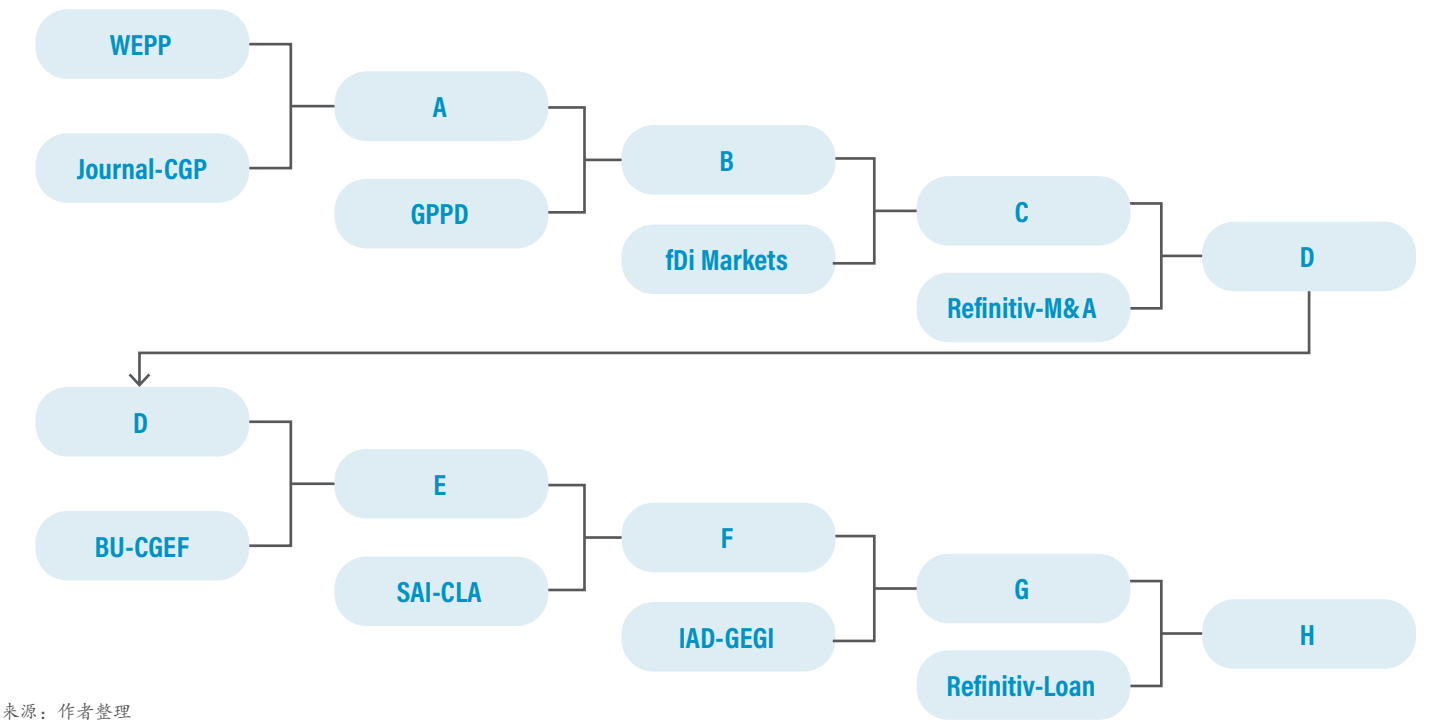
1-Jaro-Winkler距离），其取值在0（完全不相似）和1（完全匹配）之间（Cohen等人，2003）。其他模糊匹配技术用来处理数字，包括绝对差（ $|x - y|$ ）、倍数的绝对差（ $\min(|x - y|, |2x - y|, |3x - y| \dots |2y - x|, |3y - x| \dots)$ ），以及百分比差 $(x - y)/x$ 等。

图 3 | 用于匹配源数据库的技术和方法



来源：作者整理

图 2 | 匹配步骤



来源：作者整理

除定量指标外，我们还使用了“上限”和“等级”两个约束条件，提高匹配的准确性。“上限”是指返回匹配的最大距离值，例如，将Jaro-Winkler距离的上限定为0.2（即Jaro-Winkler分数为0.8），将避免匹配两个名称差异巨大的电厂。数值上限因变量而异，附录三中详细讨论匹配步骤时对此作出了进一步的说明。“等级”是指被匹配的变量之间的匹配重要性顺序。在对变量的质量和准确性有更大的信心时，变量会被赋予更高优先级，获得匹配优先权。

自动匹配和源数据所存在的限制有可能影响匹配结果的准确性。限制之一是各源数据库的项目级信息有限且不一致。一些源数据库中无可用于自动匹配的重要变量。即使是可用的变量，源数据库之间的不一致增大了成功匹配的难度。例如，由于公司名称发生变化或各源数据库在记录股权投资者时使用的公司名称不统一，各源数据库中股权投资者名称有时是不一致的。

自动匹配的另一个限制是缺乏衡量标准确保自动匹配结果的正确性。我们尝试使用一个置信度分数对可能的匹配结果进行排序，但很难通过确定一个置信度分数阈值来确保超过该阈值的结果即为正确结果。这个限制会进一步阻碍自动匹配在能源和交通部门数据库中的扩大应用，最终可能需要以人工检查的方式确保结果的准确性。

人工匹配

人工匹配按照与模糊匹配相同的顺序进行，目的是检查自动匹配的准确性并对错误进行修正。两位掌握中英双语的分析员分别对每个源数据库进行匹配，并在广泛研究和讨论的基础上消弭差异。此外，我们通过补充研究，检查原始资料的准确性，同时收集每个项目更多的相关信息，最终提高数据库的整体质量。以下为人工匹配的具体步骤：

- 信息：主要通过六类信息确定两个项目是否匹配。
 - 国家
 - 电厂燃料来源
 - 电厂名称（补充研究）
 - 投资者名称（包括债权投资者和股权投资者）（补充研究）
 - 装机容量（10%的差异）
 - 详细位置（城市和省份）
- 标准：当两个条目都在同一个国家，使用相同的燃料来源，并且上述其余四个变量中至少有三个在两个源数据库中是一致的，则认为这是一个正确的匹配。少数电厂符合上述匹配标准，但处于项目的不同阶段（一些发电项目由不同年份建造的不同电厂组成，以第一阶段、第二阶段区分），须分别记录。但大多数情况下，符合上述匹配标准的条目都是正确匹配的。在匹配时，考虑到

对容量的不同定义，允许以兆瓦（MW）为单位的容量存在大约10%的差异。

- 补充研究：通过额外的中英文资料和信息研究，确定投资者和电厂名称。每个源数据库对投资者的记录存在不同，有些使用投资公司名称，有些则使用母公司名称。
- 不匹配：未匹配成功的项目作为单独的条目保留。

两个源数据库被用来解决匹配过程中发现的两类冲突信息。不同的源数据库可能涵盖同一类型的信息，但由于收集和分类方法不同，可能造成针对同一个项目出现冲突的信息。

- 针对与电厂特征相关的冲突信息，如电厂名称和装机容量，使用WEPP为标准。WEPP提供了全球范围最全面的电厂设计信息，且采用了直接调查和二手信息收集相结合的方法（普氏，2018）。
- 针对债权投资额中出现的冲突信息，我们尽可能使用SAIS-CLA和Refinitiv-Loan贷款数据库信息为标准。这两个源数据库提供了更详细的贷款条款，并有更严谨的贷款数据收集方法论。

虽然人工匹配可以确保结果的准确性，但这一过程耗时较长，需要大量的人工投入，且很难为其他数据库和研究人员提供参考。人工匹配过程涉及使用源数据库以外的外部资源确认现有信息和寻找额外信息。这个过程并不完全客观，即使有多位审查人员，也可能存在错误。

单个投资方的贷款额估算

每个投资方的投资额信息为分析资金流提供了有用的细节。然而，这种类型的信息并不总是可以获得。所有的220个股权投资的项目都没有报告投资额。

债权投资总体上有较好的覆盖面。在253个涉及债权投资的项目中，222个报告了每个投资方的投资额（88%），其余12%的债权投资报告了所有投资方的总投资额。

我们根据源数据库中可获得的投资细节水平，从总投资中提取或估算出来自单个投资方的贷款额，具体分两种情况：

- 如果缺少单个投资方的贷款额，但同时有贷款总额和参与投资的投资方数量，则使用平均贷款额作为单个投资方的贷款额估计值。例如，BU-CGEF中的一些项目是由国家开发银行和中国进出口银行共同出资的，但未提供各自的出资额，则按照两家等额出资进行估算。
- 当源数据库中的一个条目包含了多家电厂的融资，但仅显示总投资额时，则按每家电厂装机容量加权估算其投资额。

结果

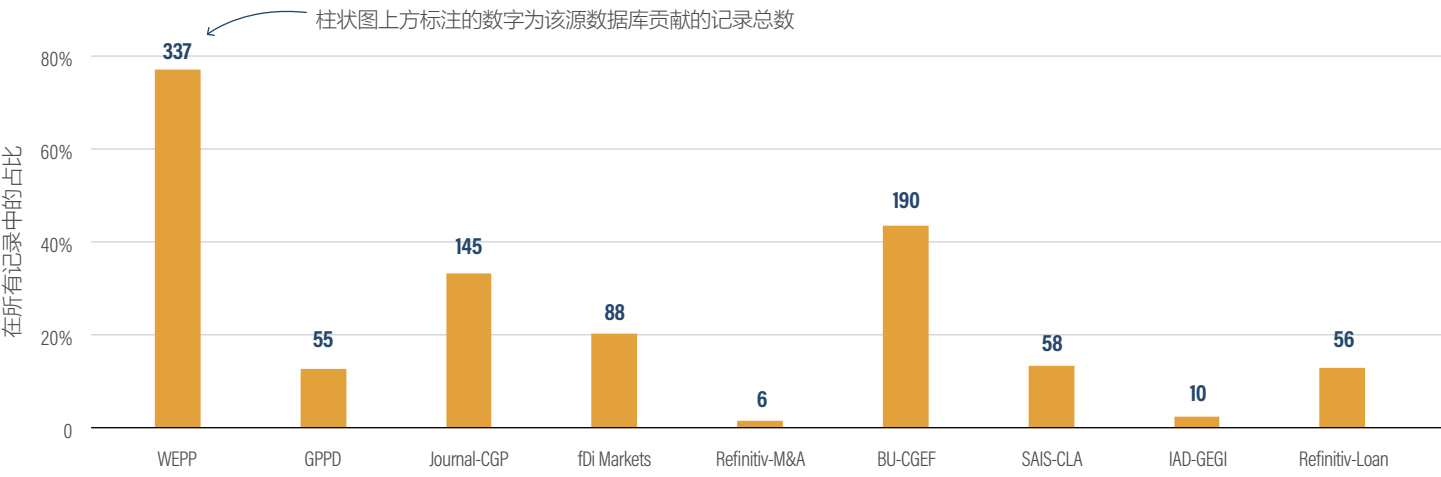
经过自动匹配和人工匹配之后，最终的数据库有443个条目，涉及430家电厂。数据库的每个条目都包括电厂的详细特征、股权和/或债权投资者的名字，以及电厂的投资额（并非全部电厂都具有此信息）。整合而成的数据库包括共计473项股权和债权投资（其中253项为债权投资，220项为股权投资）。由于有些条目同时包含了对同一个项目的债券和股权投资，数据整合后的投资数量多于条目数量。

此外，一些电厂分别在不同年份获得了多项债权投资。由于每个条目只能包括在单一年份进行的债权投资，数据库将这些交易分别记录为单独的条目。在数据库中，有两家电厂在三个不同年

份获得三次债权投资，有八家电厂在两个不同年份获得两次债权融资。因此，2家电厂有6个条目，8家电厂有16个条目，即10家电厂总共有22个条目。

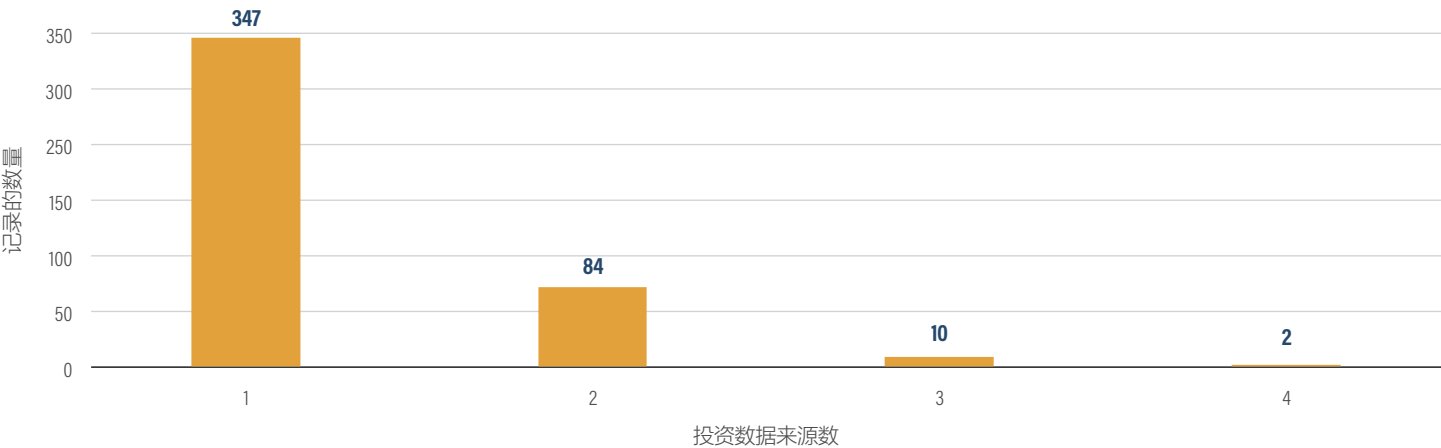
本数据库包含九个源数据库，它们的贡献各不相同，其中WEPP的贡献最大（图4）。不同的源数据库之间存在着明显的重叠。如图5所示，有占总记录的22%的96个条目来自多个投资来源。最常见的类型是有两个投资数据来源的项目。在96个项目中，有84个是有两个投资数据来源的项目。有三个或四个投资数据来源的项目相对罕见。造成各源数据库贡献差异的原因部分来自各源数据库的行业覆盖度的差异。有些源数据库覆盖了非发电项目，如SAIS-CLA就覆盖了更多交通项目。

图 4 | 数据来源对最终记录的贡献^a (N = 443)



注：a 由于一个条目结果可能有来自多个源数据库的贡献，因此这些数字的总和将大于 443。
来源：作者整理

图 5 | 记录的数量和对每个记录有贡献的投资数据来源



来源：作者整理

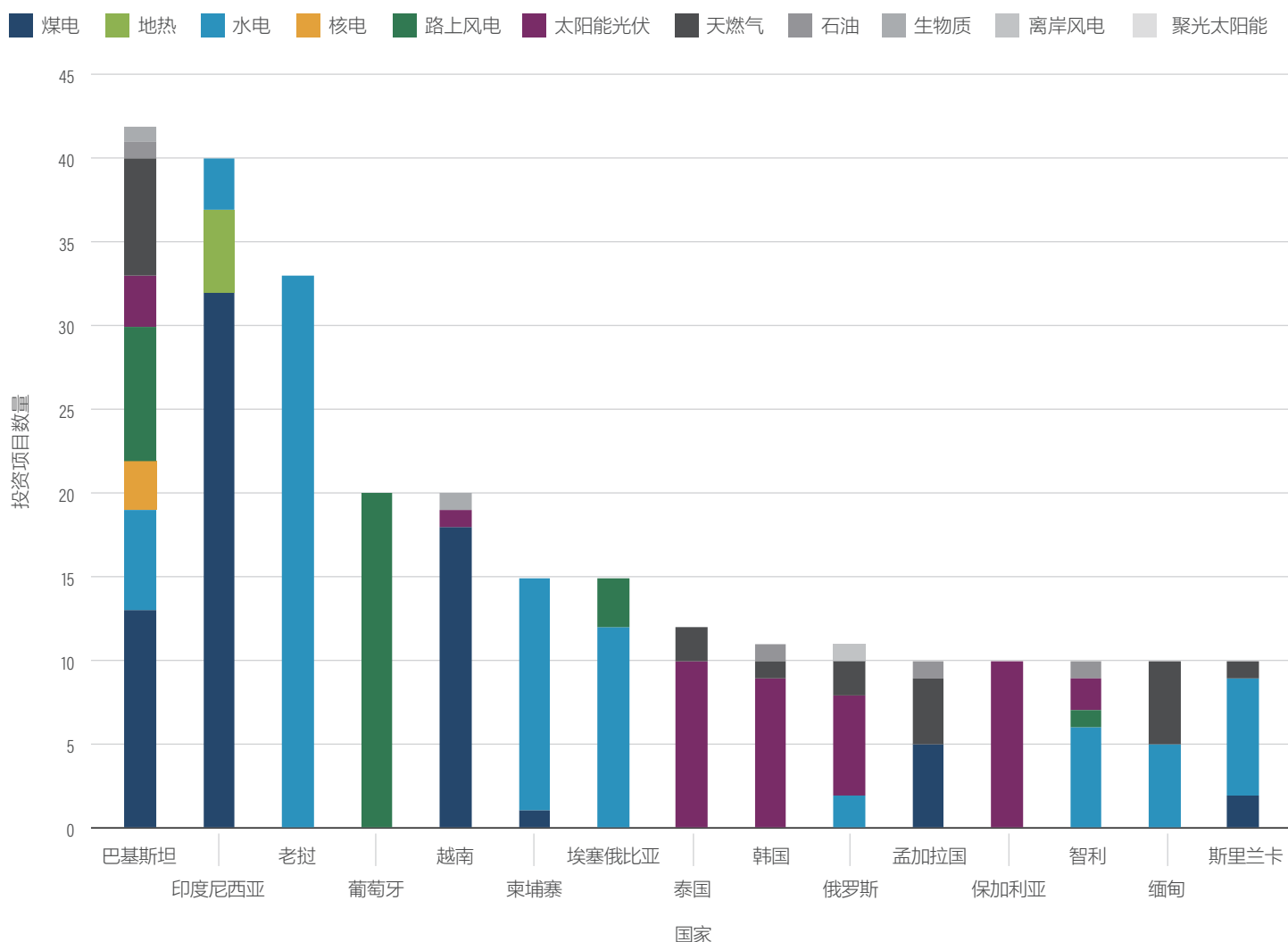
数据库中的473项投资分布在76个“一带一路”国家（图6）。按投资数量计算，排名前五的国家依次为巴基斯坦、印度尼西亚、老挝、葡萄牙和越南，共获得155项投资，约占数据库中记录的投资总数的三分之一。有15个国家获得了至少10项投资，这些国家合计获得269项投资，占投资总额的58%。数据库中共包括了253项债权投资和220项股权投资（图7）。

该数据库包括430家采用11种发电技术的电厂（图8），其中包括136家化石燃料电厂、293家非化石燃料电厂，以及1家燃料来源不明的电厂。就电厂数量而言，水电、煤电、太阳能光伏、陆上风电和天然气是投资最多的电厂类型，每种燃料类型至少有40家电厂。这五种类型的电厂共计399座，占电厂总数的93%。

由于源数据库中的股权投资额和装机容量信息缺失，数据库显示有20家电厂缺少装机容量信息，只有债权投资额可查。图10中展示了债权投资和装机容量相关数据结果，但应注意这并不反映中国海外发电投资的全貌。

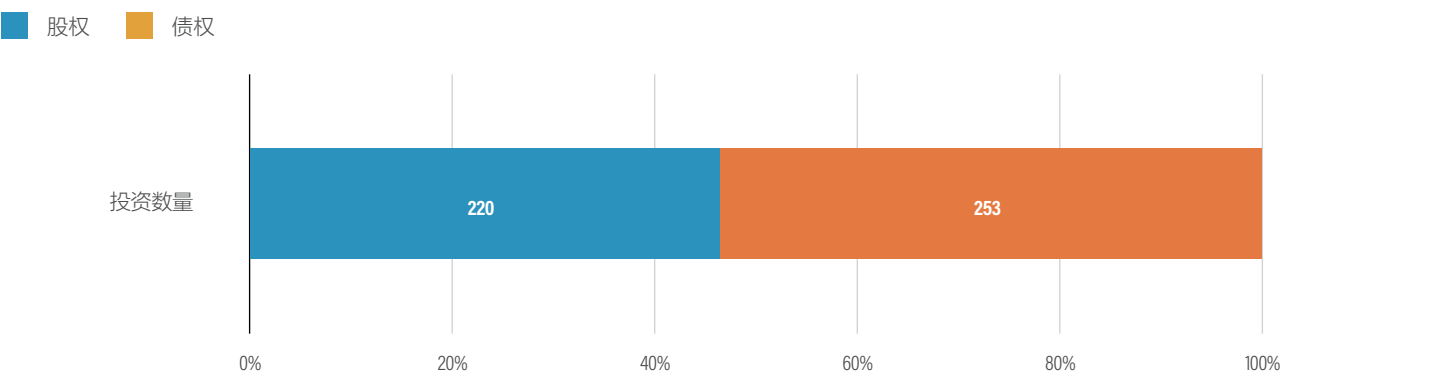
数据库结果显示，从2000年到2020年，中国投资者在“一带一路”国家总共进行了1063.5亿美元的债权投资，其中在巴基斯坦、印度尼西亚、越南、尼日利亚和老挝的投资最多（图9）。按装机容量的主要燃料分类来看，中国投资最多的是煤电部门，债权投资最多的是水电部门（图10）。

图6 | 2000年至2020年各国获得的投资项目数量



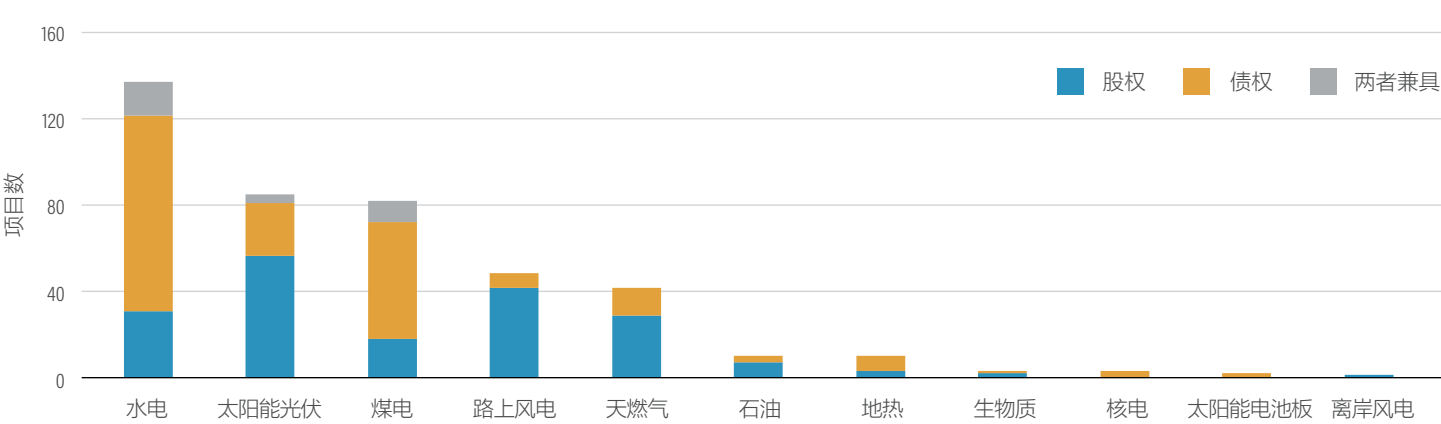
来源：作者整理

图 7 | 2000年至2020年各融资类型投资数量



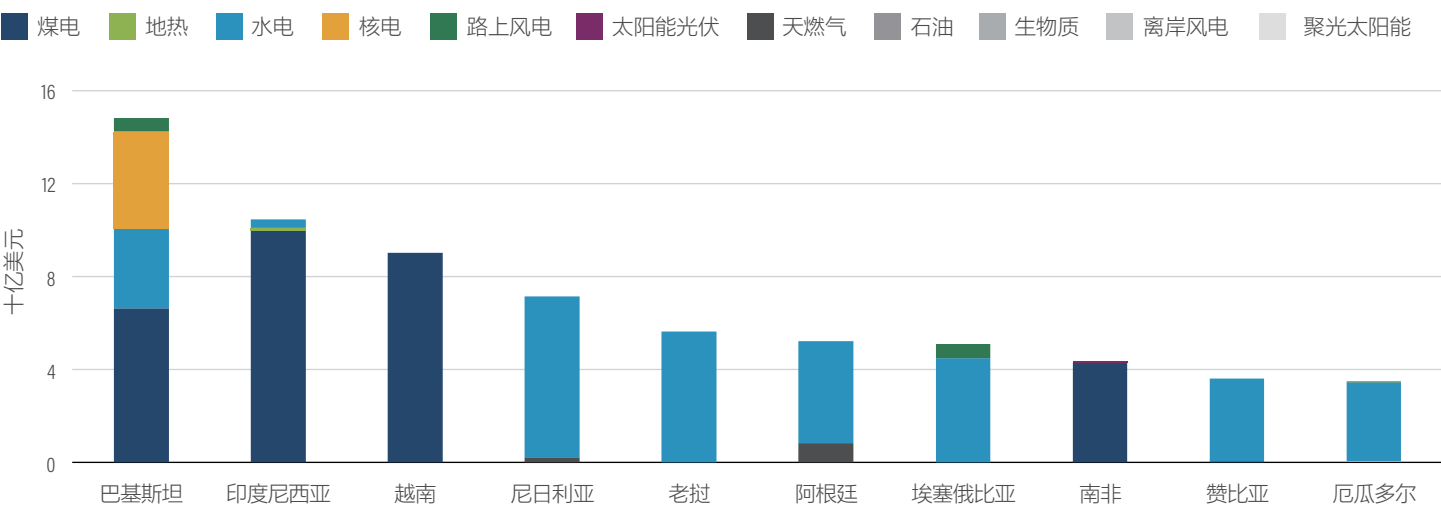
来源：作者整理

图 8 | 2000—2020年各燃料来源和投资类型的电厂数量



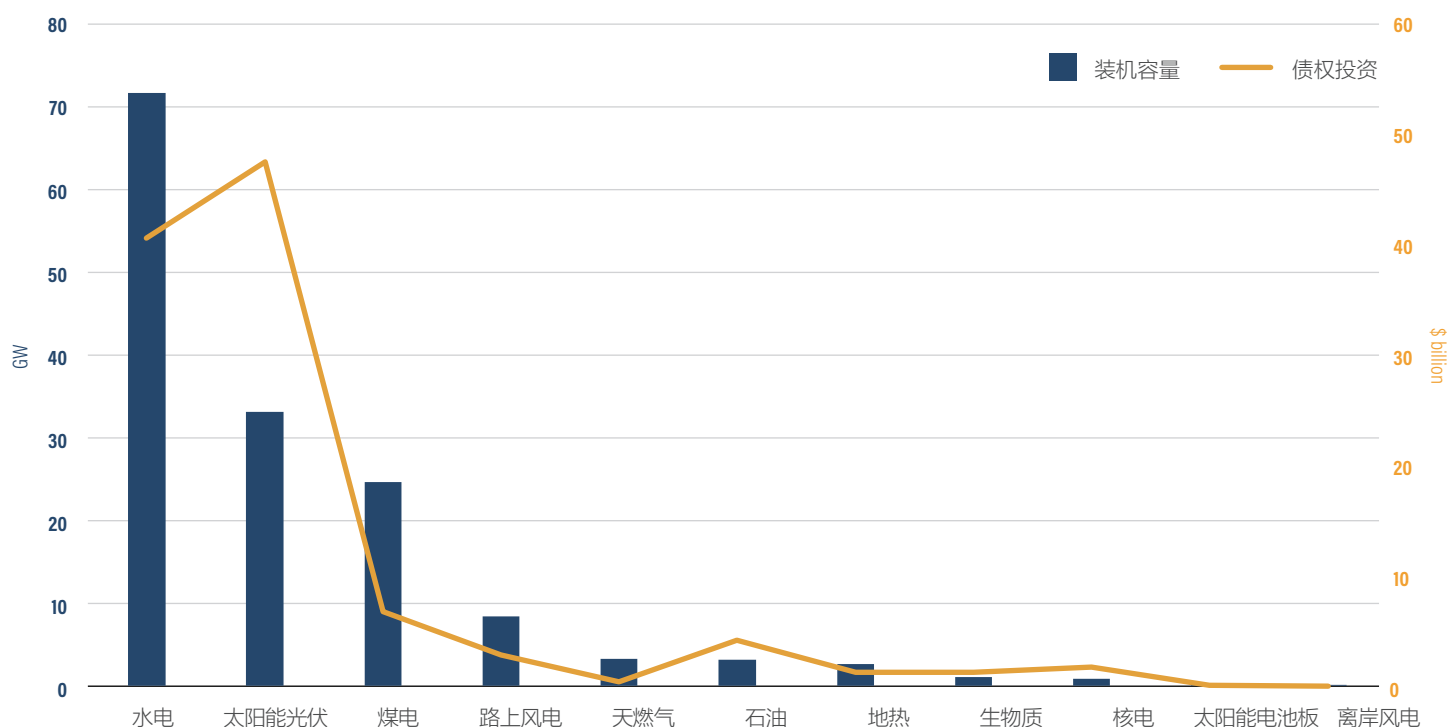
来源：作者整理

图 9 | 2000—2020年中国债权投资十大目的地投资额和各燃料来源投资额



来源：作者整理

图 10 | 2000—2020年电力投资各主要燃料装机容量和债权投资额



数据库使用和维护计划

本数据库可以为多种类型的研究提供支持和启发。例如，基于数据库对中国海外煤炭储备和债权压力进行分析，将有助于投资者管理煤炭搁浅资产风险。再如，对不同国家的石油和天然气投资以及可再生能源的趋势做出分析，可以帮助了解“一带一路”国家能源发展路径需求和差距。该数据库还可以鼓励“一带一路”国家提升数据透明度。

数据库计划每年更新，更新频率与几个源数据库一致。每年的更新将包括每个新增年份的新项目和对前几年现有项目的修正。

九个源数据库中有四个是非公开（即商业）数据库，包括WEPP、fDi Markets、Refinitiv-M&A和Refinitiv-Loan。我们已经从fDi Markets、Refinitiv-M&A和Refinitiv-Loan获得了向公众披露投资信息的许可。因产权限制，WEPP不允许将电厂设计信息公开分享，但用户可以使用该数据库中的独特标识符交叉引用WEPP中的记录。一些数据库需要每月或每年订阅，订阅过期将严重影响更新计划。

我们也在考虑扩大和改进数据库，以包括更多的投资种类和新的源数据库，并通过改进自动匹配算法和探索其他来源数据库等办法，制定解决缺失值的方案。

附录一：截至2021年1月的“一带一路”国家列表

非洲	亚洲	欧洲	大洋洲	北美洲	南美洲
苏丹	韩国	塞浦路斯	新西兰	哥斯达黎加	智利
南非	蒙古	俄罗斯	巴布亚新几内亚	巴拿马	圭亚那
塞内加尔	新加坡	奥地利	萨摩亚	萨尔瓦多	玻利维亚
塞拉利昂	东帝汶	希腊	纽埃	多米尼加	乌拉圭
科特迪瓦	马来西亚	波兰	斐济	特立尼达和多巴哥	委内瑞拉
索马里	缅甸	塞尔维亚	密克罗尼西亚联邦	安提瓜和巴布达	苏里南
喀麦隆	柬埔寨	捷克	库克群岛	多米尼克	厄瓜多尔
南苏丹	越南	保加利亚	汤加	格林纳达	秘鲁
塞舌尔	老挝	斯洛伐克	瓦努阿图	巴巴多斯	
几内亚	文莱	阿尔巴尼亚	所罗门群岛	古巴	
加纳	巴基斯坦	克罗地亚	基里巴斯	牙买加	
赞比亚	斯里兰卡	波黑			
莫桑比克	孟加拉国	黑山			
加蓬	尼泊尔	爱沙尼亚			
纳米比亚	马尔代夫	立陶宛			
毛里塔尼亚	阿联酋	斯洛文尼亚			
安哥拉	科威特	匈牙利			
吉布提	土耳其	北马其顿（原马其顿）			
埃塞俄比亚	卡塔尔	罗马尼亚			
肯尼亚	阿曼	拉脱维亚			
尼日利亚	黎巴嫩	乌克兰			
乍得	沙特阿拉伯	白俄罗斯			

非洲	亚洲	欧洲	大洋洲	北美洲	南美洲
刚果（布）	巴林	摩尔多瓦			
津巴布韦	伊朗	马耳他			
阿尔及利亚	伊拉克	葡萄牙			
坦桑尼亚	阿富汗	意大利			
布隆迪	阿塞拜疆	卢森堡			
佛得角	格鲁吉亚				
乌干达	亚美尼亚				
冈比亚	哈萨克斯坦				
多哥	吉尔吉斯斯坦				
卢旺达	塔吉克斯坦				
摩洛哥	乌兹别克斯坦				
马达加斯加	泰国				
突尼斯	印度尼西亚				
利比亚	菲律宾				
埃及	也门				
赤道几内亚					
利比里亚					
莱索托					
科摩罗					
贝宁					
马里					
尼日尔					
刚果（金）					
博茨瓦纳					

注：区域划分来源于“一带一路”网，遵循七大洲原则
来源：“一带一路”门户网站，2021

附录二：变量列表与描述

变量名称	变量描述
Unique_id	条目的唯一识别码
Power_plant_name	电厂名称
Country	电厂所在国家
Country_iso3c	联合国《国家名称用语公报》（ISO 3166-1）中的三字母国家代码
Region	联合国分类的五个地区
Subregion	联合国分类的 17 个次区域
Province	电厂所在的州或省份
City	电厂所在的城市
Installed_capacity	电厂的装机容量（MW）
Commissioning_year	电厂进入运营或计划进入运营的年份
Primary_fuel	电厂使用的主要燃料
Latitude	电厂的纬度
Longitude	电厂的经度
Duplicate_power_plant	“Y”或“N”。“Y”表示该条目中的电厂是一家重复电厂，因为同一电厂在不同年份有多项投资
Investment_type	涉及的投资类型，包括股权、债权，或两者兼有
Total_investment_estimated	“Y”或“N”。“N”表示总投资额，来自 fDi Markets 而非估计值。fDi Markets 在可变资本投资下记录这个值
Investment_cofinanced_BU	“Y”、“N”或“NA”。这是波士顿大学源数据库中的“比例未知共同出资”一栏。“Y”表示有来自国家开发银行或中国进出口银行以外的投资者的额外未知部分融资
Debt_investment_averaged	“Y”或“N”。“Y”表示单个银行的贷款额，是通过贷款总额除以贷款方总数计算的估计值。对于没有贷款方数量信息的记录，债权总额只包含了中国的部分，除以来自中国的银行的数量
Debt_investment_weighted	“Y”或“N”。“Y”表示一个债权交易对应多个发电项目，它们各自的贡献按容量加权计算
Equity_investment_weighted	“Y”或“N”。“Y”表示一个股权交易对应多个发电项目，其各自的贡献按容量加权计算
Total_investment_amount	电厂的总投资额（百万美元）
Equity_investor_name_1	第一个股权投资者的名称

变量名称	变量描述
Equity_investor_amount_1	第一个股权投资者的投资额（百万美元）
Equity_investment_type	涉及的股权投资类型
Equity_investment_year	进行股权投资的年份
Parent_company_of_investor	主要股权投资者的母公司
Debt_investment_year	进行债权投资的年份
Debt_investment_amount	电厂的债权投资总额（百万美元）
Number_of_lenders	包括非中国的银行的贷款人总数，仅在路孚特（Refinitiv）贷款数据库中存在。对于其他源数据库的贷款投资，因并未收集贷款方的总数信息，故该栏留空
Bank_of_China	中国银行的投资额（百万美元）
China_Development_Bank	国家开发银行的投资额（百万美元）
export_import_bank_of_china	中国进出口银行的投资额（百万美元）
Industrial_and_Commercial_Bank_of_China	中国工商银行的投资额（百万美元）
China_Construction_Bank	中国建设银行的投资额（百万美元）
Bank_of_Communications	交通银行的投资额（百万美元）
Agricultural_Bank_of_China	中国农业银行的投资额（百万美元）
China_Citic_Bank_Corp	中国中信银行的投资额（百万美元）
China_Merchants_Bank	招商银行的投资额（百万美元）
China_Minsheng_Banking_Corp	中国民生银行的投资额（百万美元）
China_Zheshang_Bank	中国浙商银行的投资额（百万美元）
China_Everbright_Bank	中国光大银行的投资额（百万美元）
Compid	UDI 世界电厂数据库中的“COMPID”变量，即该数据库分配给电厂母公司的唯一识别码
Unit_ID	UDI 世界电厂数据库中的“UNITID”变量，即该数据库分配给电厂发电单元的唯一识别码
Location_ID	UDI 世界电厂数据库中的“LOCATIONID”变量，即该数据库分配给电厂（地点）的唯一识别码
GPPD_ID	全球电厂数据库中的“gppd_idnr”变量，即由 10 或 12 个字符组成的电厂识别码
J_ID	《中国全球电力——中国电力行业海外投资估算》期刊数据库中普氏的“unit_id”变量

变量名称	变量描述
Fdi_id	fDi Markets 分配给一项投资的唯一识别码
RMA_ID	Refinitiv-M&A 数据库中的“交易号码”变量，即 Refinitiv 分配给一项投资的唯一识别码
BU_ID	一个由作者生成的唯一识别码，用于识别 BU-CGEF 数据库中的每个项目
SAIS_ID	中国对非洲贷款数据库中的“贷款 ID”变量，是由 SAIS 分配给贷款的唯一识别码
IAD_ID	作者生成的唯一识别码，用于识别 BU-CGEF 数据库中的每个项目
R_ID	Refinitiv 贷款数据库中的“Deal PermID”变量，即 Refinitiv 分配给一个批次贷款的唯一识别码

来源：作者整理

附录三：详细的自动匹配方案

匹配操作总共有八个步骤，我们总是将新的源数据库与之前匹配的结果源数据库进行匹配。图2给出了匹配中采取的八个步骤的概况，并将在接下来展开讨论。

表 A1 | 步骤 1: A_pp

- A_unit = WEPP + Journal-CGP at unit level
- A_pp = Sum of A_unit at power plant level

单位等值匹配	UNITID== UNITID
变量	各列的汇总标准
installed_capacity	求和（sum）
commissioning_year	取最大值（max）

由于WEPP和Journal-CGP源数据库都提供了电厂在机组层面的信息，且每个机组都有唯一的标识符“unitid”以及相同的电厂层面标识符“locationid”，我们将第一步分为两个子步骤，以便获得电厂层面的结果。首先使用WEPP和Journal-CGP的“locationid”在机组级进行等价匹配，得到源数据库“A_unit”，然后使用“locationid”和燃料类型将机组信息汇总到电厂级，得到源数据库“A_pp”。

表 A2 | 步骤 2: B

- 汇总GPPD至工厂水平B = A_pp + GPPD

等值匹配	LOCATIONID == WEPP_ID
变量	各列的汇总标准
latitude/longitude	取平均值（average）
GPPD_id	拼接（concatenated）

由于GPPD数据库中存在电厂级和机组级的混合数据，我们按“WEPP_ID”和燃料类型将数据汇总到电厂级。GPPD中的“WEPP_ID”被映射到WEPP中的“locationid”。然后，步骤B通过“locationid”和“WEPP_ID”将“A_pp”源数据库和GPPD等价匹配。

表 A3 | 步骤 3: C = B + fDi Markets

等值匹配	模糊匹配	
	等级	技术
Country, primary_fuel	equity_investor_name_n	Jaro-Winkler/cap 0.2
	commissioning_year	Absolute difference /cap 15
	installed_capacity	Absolute difference between multiples/cap 10
	city	Jaro-Winkler/ no cap
源数据库的独特贡献	变量	
B	compid, united, locationid, power_plant_name, latitude, longitude	
fDi Markets	equity_investment_year, equity_investor_amount_n, parent_company_of_investor, total_investment_amount	

第三步, 通过匹配B和fDi市场创建源数据库C。准确的匹配变量是国家和主要燃料。重要性最高的模糊匹配变量是“equity_investor_name_n”。在比较了“parent_investor_name”之后, 考虑到公司名称的准确性, 我们将其与“equity_investor_name”相匹配, 避免一些母公司拥有多个子公司, 造成名称的不匹配。

表 A4 | 步骤 4: D = C+ Refinitiv-M&A

等值匹配	模糊匹配	
	等级	技术
country	equity_investor_name_n	Jaro-Winkler / cap 0.2
	equity_investment_year	Absolute difference / cap 15 years
	parent_company_of_investor	
	equity_investment_type	
源数据库的独特贡献	变量	
C	commissioning_year, primary_fuel, installed capacity, city, province, compid, united, locationid, power_plant_name, latitude, longitude	
Refinitiv-M&A	equity_investment_type	

第四步, 通过匹配C和Refinitiv-M&A创建源数据库D。由于Refinitiv-M&A只包括并购投资, 其投资类型与直接对外投资类型（绿地投资）互不相重叠。尽管C和Refinitiv-M&A共用同一列“equity_investment_amount_n”, 但这个变量只出现在fDi Markets的源数据库中, 因此源数据库C中的金额只包括绿地投资的金额。我们将“equity_investment_amount_n”作为附加信息。

等值匹配使用的变量是国家。但Refinitiv-M&A所记录的交易存在以下问题: 这些交易是在公司层面上发生的, 通常只显示一家中国公司收购或兼并外国的另一家公司, 而没有关于该公司拥有的电厂类型或数量的详细信息。在大多数情况下, 该公司拥有一个以上的电厂, 使用不同的主要燃料（如风能和煤）。这给我们与其他以电厂为观测单位的源数据库进行匹配带来了困难, 大大降低了成功匹配的概率。这个问题通过人工匹配得到了解决。

债权投资

表 A5 | 步骤 5: E = D + BU-CGEF

等值匹配	模糊匹配	
	等级	技术
Country, primary_fuel	power_plant_name	Jaro-Winkler / cap 0.2
	installed_capacity	Absolute difference between multiples / cap 300 mw
源数据库的独特贡献	变量	
D	compid, united, locationid, latitude, longitude, equity_investment_type, equity_investor_name_n, equity_investment_amount_n, equity_investment_year, parent_company_of_investor	
BU-CGEF	debt_investment_year, china_development_bank, export-import_bank_of_china	

第五步, 向数据库添加债权投资信息。这一步主要增加了国家开发银行和中国进出口银行的债权投资信息, 并确定了同时具有股权和债权投资的项目。

表 A6 | 步骤 6: F = E + SAIS-CLA

等值匹配	模糊匹配	
	等级	技术
country	debt_investment_year	Absolute difference / cap 15 years
	power_plant_name	Jaro-Winkler / cap 10
	installed_capacity	Absolute difference between multiples / cap 300 mw
	Debt investment amounts (china_development_bank, export-import_bank_of_china)	Percent difference /cap 10%
源数据库的独特贡献	变量	
E	compid, united, locationid, latitude, longitude, province, city, primary_fuel, equity_investment_type, equity_investor_name_n, equity_investment_amount_n, equity_investment_year, parent_company_of_investor	
SAIS-CLA	industrial_and_commercial_bank_of_china, china_construction_bank, bank_of_china, agricultural_bank_of_china	

由于 BU-CGEF数据库只有国家开发银行和中国进出口银行的债权投资额数据，所以我们单独匹配了关于国家开发银行和中国进出口银行债权投资额的数据。考虑到投资额记录的不确定性，我们把债权投资额放在层次结构的最后。有时，一些共同融资项目的贷款方被合并，仅总投资额可用。在这种情况下，我们依靠电厂名称和装机容量进行匹配。

表 A7 | 步骤 7: G = F + IAD-GEGI

等值匹配	模糊匹配	
	等级	技术
Country	debt_investment_year	Absolute difference / cap 15 years
	Debt investment amounts (all Chinese banks)	Percent difference /cap 10%

IAD-GEGI源数据库和F源数据库共享债权投资银行列，匹配技术与上一步相同。截至第七步，所有必要的栏目都已建立，因此IAD-GEGI的作用为添加新的交易记录和补充匹配后缺失的数据值。

表 A8 | 步骤 8: H = G + Refinitiv Loan

等值匹配	模糊匹配	
	等级	技术
Country	debt_investment_year	Absolute difference / cap 15 years
	power_plant_name	Jaro Winkler
	Debt investment amounts （ all Chinese banks ）	Percent difference /cap 10%

最后一步是加入源数据库Refinitiv Loan。该源数据库仅记录债权投资，但其中单个债权投资额大多不完整，因此基于这一变量的匹配可信度较低。为此，我们从该源数据库的描述栏目中提取了电厂名称。

在完成所有的匹配步骤后，我们最终确定了结果。首先，我们删除了没有投资数据的行，并计算每笔交易的总投资额。然后，根据识别码更新了投资类型。例如，如果一个项目所有债权投资来源的识别码都是“NA”，那么该项目的融资类型为股权。如果一个项目没有一个识别码是“NA”，那么该项目的融资类型为债权与股权兼具。最后，我们根据最终数据变量列表对该列进行排序，完成最终呈现。

尾注

BRI	“一带一路”倡议
BU-CGEF	波士顿大学中国全球能源融资数据库
CARI	中国—非洲研究倡议
China Eximbank	中国进出口银行
FDI	直接对外投资
GPPD	全球电厂数据库
IAD	美洲对话
M&A	兼并与收购
SAIS-CLA	国际高级研究学院—中国对非洲贷款
WEPP	世界电厂数据库

尾注

1. 我们将对外投资绿地投资和并购作为独立的股权投资进行分析，因为它们代表了不同的投资策略，对中国投资者意味着显著不同的风险水平。对外投资绿地投资因涉及在海外建立新的实体，较并购的风险更大。
2. 资产负债表融资通常是由金融机构直接向母公司提供贷款或信贷额度，而不是为基础设施项目设立的特殊目的机构提供融资。公司债券是由公司发行的债券，面向更广泛的投资者筹集资金。担保债务凭证是一种资产支持证券，持有如银行贷款或公司债券等不同种类的债务，出售给投资者。外国中央银行可以使用中国人民银行的互换额度借入人民币，主要用于解决短期流动性问题和贸易结算（新华社，2018）。
3. 签署贷款合同通常表发电项目融资协议得到实际执行，即确认融资额度将提供给该项目。
4. 整合的数据库中不包括独立的输配电项目，但包括那些以联网为目的的发电项目。以联网为目的的发电项目在数据库中与其他电厂项目合并。自备电厂是为用户提供局部电力的电厂，通常为工业设施（Clarke Energy, n.d.）。
5. 当一个项目的电厂名称、容量和位置三条信息中有两条缺失时，则被认为是具体情况缺失的项目。
6. 如果电厂特征信息和投资信息相同，则认为项目是重复的。所有重复的项目都在路孚特（Refinitiv）贷款数据库中。该数据库重复的贷款项目是因为将有多个贷款档分期的项目重复计入。
7. 除了与数据库中的变量直接相关的信息，我们还提取了一些有助于组织和整理源数据库的额外变量，特别是从Refinitiv和Refinitiv-M&A中提取的变量。我们从Refinitiv中提取了融资类型的信息，以帮助我们更好地定位项目融资条目；我们从Refinitiv-M&A中提取了交易类型的信息，以更好地了解一项交易是收购一家公司、一家电厂，还是收购多家电厂。在这两个数据集，我们提取了能源子行业的信息，对项目是发电、输电还是配电进行分类。
8. 例如，“铭牌容量”是指发电机、原动机或其他电力生产设备在制造商指定的特定条件下的最大额定输出。“净容量”是指从铭牌容量中减去电厂的寄生负载后的电厂容量（USEIA, n.d.; NREL, 2013）。
9. fDi Markets不提供股权投资额，只包括资本投资，即项目的总投资成本。如果没有股权比例等额外信息，则无法计算出股权投资额。
10. 平均贷款出资额是将总贷款额除以参与方数量得出。在少数情况下，源数据中有在所有贷款方中一到两个贷款方的出资额。我们在处理数据时，在总金额中减去已知的来自贷款方的金额，将剩余的金额平均分配给所有其他贷款方。
11. 成本超支、初始投资不足以完成项目，项目需要额外的投资时，则会出现此类情况。
12. 我们在数据库中创建了新变量，用于记录针对同一电厂的额外条目。

参考文献

- Arel-Bundock, V., C.J. Yetman, and N. Enevoldsen. 2020. *Countrycode: Convert Country Names and Country Codes* (version 1.2.0). <https://CRAN.R-project.org/package=countrycode>. Accessed April 20, 2021.
- Belt and Road Portal. 2021. “已同中国签订共建‘一带一路’合作文件的国家一览 [Countries That Have Signed Co-Developing Belt and Road Collaboration Documents with China].” 2021. <https://www.yidaiyilu.gov.cn/gbjg/gbgk/77073.htm>. Accessed March 31, 2021.
- Brautigam, D., and J. Hwang. 2020. “SAIS-CARI Research Guidebook.” SAIS China-Africa Research Initiative. <https://static1.squarespace.com/static/5652847de4b033f56d2bdc29/t/5efd04572cbaa4329d32be1e/1593640026011/SAIS-CARI+Research+Guidebook+2020.07.01.pdf>.
- BRIGC. n.d. “BRIGC Overview.” http://en.brigc.net/About_us/Overview/202009/t20200928_102502.html. Accessed May 10, 2021.
- Byers, L., J. Friedrich, R. Hennig, A. Kressig, X. Li, L.M. Valeri, and C. McCormick. 2018. “A Global Database of Power Plants,” May. <https://www.wri.org/research/global-database-power-plants>.
- CARI (China Africa Research Initiative), Boston University Global Development, and Boston University Global Development Policy Center. 2021. (Database.) *Chinese Loans to Africa Database, Version 2.0*. 2021. <https://chinaafricaaloandata.org/>. Accessed March 31, 2021.
- Clarke Energy. n.d. “Captive Power Plant.” <https://www.clarke-energy.com/applications/captive-power-plants/>. Accessed April 19, 2021.
- Cohen, W., P. Ravikumar, and S. Fienberg. 2003. “A Comparison of String Metrics for Matching Names and Records.” *Kdd Workshop on Data Cleaning and Object Consolidation* 3: 73–78.
- Dentons. 2013. “A Guide to Project Finance.” Dentons. <https://www.dentons.com/-/media/6a199894417f4877adea73a76caac1a5.ashx>. Accessed May 12, 2021.
- DOE (U.S. Department of Energy). 2015. “The Five-Step Process Framework for Project Development.” <https://www.energy.gov/sites/default/files/2015/08/f25/5-Step%20Project%20Development%20Overview.pdf>.
- Financial Times. 2021. “FDi Markets.” 2021. <https://www.fdimarkets.com/>. Accessed March 31, 2021.
- Gallagher, K.P. 2021. “China’s Global Energy Finance.” <http://www.bu.edu/cgef/#/intro>. Accessed March 31, 2021.
- Gallagher, K.P., and M. Myers. 2021. (Database.) *China-Latin America Finance Database*. https://www.thedialogue.org/map_list/. Accessed March 31, 2021.
- Global Energy Observatory, Google, KTH Royal Institute of Technology in Stockholm, Enipedia, and World Resources Institute. 2018. (Database.) *Global Power Plant Database Version 1.2.0*. <https://datasets.wri.org/dataset/globalpowerplant-database>. Accessed March 31, 2021.
- Hilman, J., and A. Tippet. 2021. “The Climate Challenge and China’s Belt and Road Initiative.” Council on Foreign Relations. Blog. March 31, 2021. <https://www.cfr.org/blog/climate-challenge-and-chinas-belt-and-road-initiative>.
- IEA (International Energy Agency). 2020. “Energy Financing and Funding—World Energy Investment 2020—Analysis.” IEA. <https://www.iea.org/reports/world-energy-investment-2020/energy-financing-and-funding>. Accessed August 30, 2021.
- Li, Z., K.P. Gallagher, and D.L. Mauzerall. 2020. “China’s Global Power: Estimating Chinese Foreign Direct Investment in the Electric Power Sector.” *Energy Policy* 136 (January): 111056. doi:10.1016/j.enpol.2019.111056.
- Mawutor, J.K.M. 2015. “Role of Project Finance in Emerging Economies.” SSRN Scholarly Paper ID 2258834. Rochester, NY: Social Science Research Network. doi:10.2139/ssrn.2258834.
- MFA (Ministry of Foreign Affairs). 2015. “Vision and Actions on Jointly Building Silk Road Economic Belt and 21st-Century Maritime Silk Road.” https://www.fmprc.gov.cn/mfa_eng/zxxx_662805/t1249618.shtml. Accessed May 10, 2021.
- NREL (National Renewable Energy Laboratory). 2013. “Solar Energy and Capacity Value.” NREL. <https://www.nrel.gov/docs/fy13osti/57582.pdf>.
- Platts. 2018. (Database.) *UDI World Electric Power Plants Database*. S&P Global. <https://www.spglobal.com/platts/en/products-services/electric-power/market-data-power>. Accessed March 31, 2021.
- R Core Team. 2020. (Database.) *R: A Language and Environment for Statistical Computing (version 4.0.3)*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>. Accessed April 22, 2021.
- Refinitiv. 2021. Refinitiv Eikon. <https://www.refinitiv.com/en/products/eikon-trading-software>. Accessed March 31, 2021.
- Robinson, D., J. Bryan, and J. Elias. 2020. (Database.) *Fuzzyjoin: Join Tables Together on Inexact Matching* (version 0.1.6). <https://CRAN.R-project.org/package=fuzzyjoin>. Accessed April 22, 2021.
- Sato, I., B. Elliott, and C. Schumer. 2021. “What Is Carbon Lock-in and How Can We Avoid It?” <https://www.wri.org/insights/carbon-lock-in-definition>. Accessed August 23, 2021.
- US EIA (U.S. Energy Information Administration). n.d. “Glossary—U.S. Energy Information Administration (EIA).” <https://www.eia.gov/tools/glossary/index.php>. Accessed April 26, 2021.
- Wang, C.N. 2021. *China Belt and Road Initiative (BRI) Investment Report H1 2021—Green Belt and Road Initiative Center*. <https://green-bri.org/china-belt-and-road-initiative-bri-investment-report-h1-2021/>. Accessed September 10, 2021.
- Xinhua News Agency. 2018. “中巴两国央行续签双边本币互换协议 [The Central Banks of China and Pakistan Renew the Bilateral Local Currency Swap Agreement].” http://www.gov.cn.qingcdn.com/xinwen/2018-05/24/content_5293454.htm.
- Zhou, L., S. Gilbert, Y. Wang, M.M. Cabré, and K.P. Gallagher. 2018. *Moving the Green Belt and Road Initiative: From Words to Actions*. Washington, DC: World Resources Institute. <https://files.wri.org/s3fs-public/moving-green-belt-and-road-initiative-from-words-to-actions.pdf>.

致谢

感谢ClimateWorks基金会对我们工作的支持和资助。感谢以下同事作为内部审核专家在写作过程中提出的宝贵意见：Dimitrios Mentis、Laura Malaguzzi Valeri、Logan Byers、王珮珊、王烨、刘哲等。感谢以下专家作为外部审核专家给与的支持：Christoph Nedopil、Cecilia Han Springer、徐佳君、刘栩等。同时，感谢来自 Sonoma Technology公司的Nathan Pavlovic和Cari Gostic，因为他们在机器学习方面的专业知识和对本项目数据自动匹配方面的支持，这篇技术说明才得以完成。尽管非常感谢审稿人的贡献，但本文中的观点仅反映了作者的观点。我们还要感谢Renee Pineda和世界资源研究所研究与开发团队在出版过程中给予的支持，感谢他们付出的时间和努力；感谢Caroline Taylor、Romain Warnault和Billie Kanfer的编辑、设计和润色。

关于作者

周李焕是世界资源研究所金融中心的研究员。邮箱：lihuan.zhou@wri.org

马子轶是世界资源研究所金融中心的助理研究员。邮箱：ziyi.ma@wri.org

刘爽是世界资源研究所的中国金融高级研究员和金融中心代理全球主任。邮箱：shuang.liu@wri.org

马欣悦是波士顿大学全球发展政策研究中心的中国研究项目总监。邮箱：xma@bu.edu

Kevin Acker曾是约翰霍普金斯大学国际关系高级学院中非研究所的研究员。

Margaret Myers是美洲对话亚洲和拉美项目的主任。邮箱：mmyers@thedialogue.org

Deborah Brautigam博士是Bernard L. Schwartz政治经济学教授和约翰霍普金斯大学国际关系高级学院中非研究所的主任。邮箱：dbrautigam@jhu.edu

Kevin P. Gallagher是波士顿大学帕迪全球研究学院的教授以及波士顿大学全球发展政策研究中心主任。邮箱：kpg@bu.edu

关于约翰霍普金斯大学中非研究所

约翰霍普金斯大学中非研究所于2014年启动，总部设在华盛顿特区的约翰-霍普金斯大学高级国际研究学院，研究所的建立是为了通过高质量的数据收集、实地研究、会议和合作，促进对中国和非洲国家关系的循证理解。我们的使命是促进研究，进行基于证据的分析，促进合作，并培训未来的领导人，以更好地了解中国和非洲关系的经济和政治层面及其对人类安全和全球发展的影响。

关于美洲对话组织

美洲对话组织汇集各国领袖，并推动拉丁美洲区域内的民主治理、社会平等和经济发展。组织与领袖努力塑造政策辩论，设计解决方案，共同商讨解决西半球的问题和挑战。

关于波士顿大学全球发展政策研究中心

波士顿大学全球发展政策研究中心是一个一政策为导向的研究中心，致力于推动全球的金融稳定，人类福祉和环境可持续性。了解更多：bu.edu/gdp。